Oikos 123: 1250–1260, 2014
doi: 10.1111/oik.00857

Subject Editor: Rampal S. Etienne. Accepted 25 February 2014

Detecting phylogenetic signal in mutualistic interaction networks using a Markov process model

H. O. Minoarivelo, C. Hui, J. S. Terblanche, S. L. Kosakovsky Pond and K. Scheffler

H. O. Minoarivelo and K. Scheffler (kscheffler@ucsd.edu), Computer Science Division, Dept of Mathematical Sciences, Stellenbosch Univ., Matieland 7602, South Africa. – HOM and C. Hui, Centre for Invasion Biology, Dept of Mathematical Sciences, Stellenbosch Univ., Matieland 7602, South Africa. – KS and S. L. Kosakovsky Pond, Dept of Medicine, Univ. of California, San Diego, USA. – J. S. Terblanche, Centre for Invasion Biology, Dept of Conservation Ecology and Entomology, Stellenbosch Univ., Matieland 7602, South Africa.

Ecological interaction networks, such as those describing the mutualistic interactions between plants and their pollinators or between plants and their frugivores, exhibit non-random structural properties that cannot be explained by simple models of network formation. One factor affecting the formation and eventual structure of such a network is its evolutionary history. We argue that this, in many cases, is closely linked to the evolutionary histories of the species involved in the interactions. Indeed, empirical studies of interaction networks along with the phylogenies of the interacting species have demonstrated significant associations between phylogeny and network structure. To date, however, no generative model explaining the way in which the evolution of individual species affects the evolution of interaction networks has been proposed. We present a model describing the evolution of pairwise interactions as a branching Markov process, drawing on phylogenetic models of molecular evolution. Using knowledge of the phylogenies of the interacting species, our model yielded a significantly better fit to 21% of a set of plant–pollinator and plant–frugivore mutualistic networks. This highlights the importance, in a substantial minority of cases, of inheritance of interaction patterns without excluding the potential role of ecological novelties in forming the current network architecture. We suggest that our model can be used as a null model for controlling evolutionary signals when evaluating the role of other factors in shaping the emergence of ecological networks.

Ecological networks are a powerful tool for representing the species interactions of a complicated ecosystem (Bascompte and Jordano 2007, Hui et al. 2013). Of special interest are mutualistic networks formed by the reciprocal dependence of plants on their pollinators, frugivores or seed dispersers. This mutualistic dependence can lead to complicated co-evolutionary dynamics (Rodríguez-Gironés and Llandres 2008, Zhang et al. 2013) and further contributes to weaving a complex web, affecting how ecosystems function and how stability is retained under anthropogenic or environmental perturbations (Kaiser-Bunbury et al. 2010). Understanding the processes that form and sustain a mutualistic network is, therefore, of pivotal importance for better managing natural resources and predicting the impacts of perturbations such as biodiversity loss, biological invasion, climate change, or habitat transformation on ecosystem stability and function.

Mutualistic networks typically exhibit several key patterns. First, early studies on the level of species generalism and specificity in food webs (Waser et al. 1996, Vázquez and Aizen 2003) motivated the examination of how the interaction degree of a species (i.e. the number of species in the network with which it interacts) is distributed. It was found that most species are poorly connected, with only a small number being well connected, which results in a right-skewed degree distribution. In most cases, the degree distribution follows a truncated power law (Jordano et al. 2003; but see Okuyama 2008) while in others it follows either a power law distribution or an exponential distribution. Second, a mutualistic network is often nested, meaning that specialists interact with species that form subsets of the species with which generalists interact (Bascompte and Jordano 2007). Bascompte et al. (2006) further suggest that this nested structure is often highly asymmetric: plant species may depend strongly on animal species but not necessarily the reverse. Another important feature of mutualistic networks is the existence of modules (also called compartments) where species strongly interact almost exclusively with species in the same module (Dicks et al. 2002). Evidently, these multiple features of mutualistic networks are not independent of each other, suggesting that an integrated model is required to better capture the intrinsic dynamic features of species interactions.

Significant progress has been made in proposing plausible models to explain some or most of the above patterns. While some studies explain the observed structure

by a direct exploration of datasets, others build models that incorporate processes and mechanisms that could be responsible for specific network structures. In the former case, a 'neutral' hypothesis has been proposed (Vázquez 2005): the species interaction in a network reflects random encounter of individuals and thus depends only on the relative abundance of species in the network. This neutral hypothesis has been suggested as a potential explanation for asymmetrical interactions, such as the nested structure in some networks (Vázquez et al. 2007, Krishna et al. 2008). In addition, the spatiotemporal variation of species distribution and the resultant sampling artifacts can also account for the presence or absence of certain interactions and some structures of mutualistic networks (Olesen et al. 2008, Vázquez et al. 2009). Indeed, Morales and Vázquez (2008) showed that the movement of frugivorous birds can change the interaction probabilities with fruit plants and thus the network structures. Temporal overlap of phenology between interacting partners can also affect the interaction strength and thus the robustness of mutualistic networks (Encinas-Viso et al. 2012).

In the latter case, some mechanistic models have been developed based on Barabási and Albert's (1999) growth and preferential attachment rules (Vázquez 2005, Takemoto and Arita 2010). Scale free networks could arise through the process of preferential attachment, whereby a new species coming into the community is more likely to interact with a generalist (a high degree node) rather than a specialist. For instance, Vázquez (2005) suggested that the truncated power law of the degree distribution could be expected if the node degree and the frequency of interactions are correlated. Guimarães et al. (2007) used a differential growth rate of node degrees and derived a truncated power law node-degree distribution. The high level of nestedness typical of interaction networks has also been shown to emerge from the preferential attachment rule (Medan et al. 2007) or from the optimization of interactions to ensure the increase of population size at equilibrium (Suweis et al. 2013). Other models have been built in the Lotka–Volterra framework (Holland and DeAngelis 2010, Zhang et al. 2011). For instance, Zhang et al. (2011) introduced both a functional response and an adaptive interaction switch, which allows re-wiring between interacting species, into a Lotka–Volterra model where multiple patterns emerge simultaneously. This allowed them to successfully predict measurements of these patterns in 81 empirical networks. Finally, some models define the occurrence of an interaction based on matching morphological and behavioural traits (i.e. trait complementarity and exploitation barriers). For example, the modularity in mutualistic networks is often associated with morphological traits (Olsen et al. 2007), where coevolution favours trait complementarity and convergence (Nuismer et al. 1999, Guimarães et al. 2011).

Mutualistic interactions are affected by species phenotypes, which in turn reflect the evolutionary history of the species. Interactions may therefore also be influenced by phylogenetic constraints. Thus a group of ontogenetic models – on which we build the present work – incorporates the evolutionary history (phylogeny) of the interactive taxa. Specifically, Rezende et al. (2007a) simulated the evolution of phenotypic traits along phylogenetic trees, which partially explained the nested architecture in mutualistic networks. Rezende et al. (2007b) further found that, in about half the cases, the phylogenetic distances between paired species are positively correlated with their interaction similarities, suggesting a tendency for closely related species to interact with the same partners. Although the phylogenetic signal in mutualistic networks is evident, the function of the evolutionary process in constructing the network architecture remains unclear. Only a few studies have explored the way in which the evolution of individual species affects the dynamics of interaction networks. These include studies hypothesizing selective effects: Guimarães et al. (2011), who modelled the co-evolution of traits in mutualistic networks and found a higher convergence of traits in super-generalist species, which in turn plays an important role in the maintenance of the community organization and stability. Nuismer et al. (2013) showed that co-evolutionary selection may increase the network connectance while altering the pattern of nestedness. To our knowledge, however, no generative explanation of the way in which the evolution of individual species affects the dynamics of interaction networks has been proposed, even in the absence of selective effects.

Here, we present a model describing the evolution of pairwise interactions as Markov processes, drawing on phylogenetic models of molecular evolution (Felsenstein 2004). Our model does not require any assumption of selective effects, and is based on the assumption that individuals inherit not only the genetics of their parents, but also their interaction partners. When interaction patterns persist over evolutionary time scales, as can be expected particularly when interactions are closely linked to morphology and thus to underlying genotypes, this also leads to species inheriting the interaction patterns of their ancestors. Using this model, we systematically explore three questions using 53 empirical networks. First, we determine whether known interaction networks are better explained by real phylogenetic trees than by randomly generated phylogenetic trees. This allows us to evaluate the overall importance of phylogenetic history in the evolution of the currently observed network. Second, we examine whether networks simulated under our model provide a better fit to the observed node-degree distribution than do other contending models, and whether the nestedness and modularity properties predicted by our model are consistent with empirically observed values. Third, we explore the inferred gain–loss rates of specific interactions along the evolutionary history; that is, we analyse the average heritability of interactions over evolutionary time.

Methods

Model overview

Our model is based on an approach that has become standard in phylogenetic modelling: see Durbin et al. (1999), Felsenstein (2004), Yang (2006) and Lemey et al. (2009) for introductions to this discipline. Our description proceeds in three steps: 1) in 'Markov process of interaction evolution', we define a Markov process describing the appearance and

disappearance of interactions between each pair of species over evolutionary time; 2) in 'Constructing an evolutionary tree of species pairs', we describe the construction of a tree in which each pair of contemporaneous ancestral species is represented as a branch; 3) finally, in 'Network evolution on the combined tree', we apply the above Markov process to every branch in the tree, resulting in a single Markov process that describes the evolution of an entire interaction network. Steps 1 and 3 are standard in phylogenetic modelling, but step 2 is a novel extension specific to the evolution of interaction networks.

Markov process of interaction evolution

We consider the evolution of the interaction between a particular pair of species as a discrete state, homogeneous continuous-time Markov process for a binary character (Churchill 1989, Stroock 2005). The process describes the probability of changing between two possible states: the absence of an interaction (state s_0), and the presence of an interaction (state s_1), over time ($t \geq 0$). Following the parameterization that has become standard in phylogenetic modelling, we define the time-homogeneous Markov process by its instantaneous rate matrix $\boldsymbol{Q}$:

$$\begin{array}{c} \\ s_0 \\ s_1 \end{array} \begin{array}{c} \begin{array}{cc} s_0 & s_1 \end{array} \\ \begin{bmatrix} -\mu\pi_1 & \mu\pi_1 \\ \mu\pi_0 & -\mu\pi_0 \end{bmatrix} \end{array}$$

Here, μ is a parameter controlling the rate of interaction gain (or loss), measured in interaction changes per unit of branch length; π_0 and π_1 ($\pi_0 + \pi_1 = 1$) are parameters determining the equilibrium frequencies of s_0 and s_1, respectively. For example, the probability of finding the process in state s_0 is π_0 as time approaches infinity, regardless of the starting state. The process is assumed to be at equilibrium throughout, with the initial state drawn from $\{s_0, s_1\}$ with probabilities $\{\pi_0, \pi_1\}$. The corresponding transition probability matrix $P(t) = e^{Qt}$ contains the probabilities of changing from state s_i to s_j ($i, j = 0$ or 1) within an amount of time $t \geq 0$.

Constructing an evolutionary tree of species pairs

Next, we consider the problem of reconstructing the evolutionary history of an observed bipartite ecological network, where each interaction is between members of two disjoint sets of species (e.g. plants and pollinators), with a dated phylogeny given for each set. Each potential interaction in the observed network corresponds to an extant 'species pair', and each potential historical interaction corresponds to an ancestral 'species pair'. We therefore use the phylogenies to reconstruct, at each point in time, the pairs of ancestral species that co-existed. Whenever a speciation event occurs in either of the phylogenetic trees, the result is that an ancestral species pair is replaced with two new species pairs: this leads to the key insight that, just as the evolutionary splitting of individual species can be represented as a tree structure (referred to as a phylogeny), the evolutionary splitting of 'species pairs' can also be represented as a tree structure, which we shall call the 'combined tree' (Fig. 1). The leaf nodes of the combined tree represent the species pairs (whether interacting or not) in the observed interaction network and, going back in time, each internal node represents a particular ancestral species pair at a point in time that is known from the dated phylogenies.

To construct the combined tree (Fig. 1), we start with the oldest branch (Br 1A) representing the combination of the pair of oldest branches of the two phylogenetic trees (Br 1 and Br A). Next, the split at node 1 in phylogenetic tree 1 is represented in the combined tree by the split of the branch Br 1A into Br 2A and Br 3A; going forward in time, the split at node A in phylogenetic tree 2 is represented by splitting Br 2A into Br 2B and Br 2C and, simultaneously, Br 3A into Br 3B and Br 3C. We continue this procedure until the leaf nodes are reached.

Although phylogenies were available for each set of species in our interaction networks, the branch lengths were unknown in most cases, and estimating them from the available network data would not have been statistically feasible. In order to obtain dated phylogenies from which to construct the combined tree, we approximated the branch lengths using a tree shape parameter λ (corresponding to ρ as introduced by Grafen 1989) and a relative scale

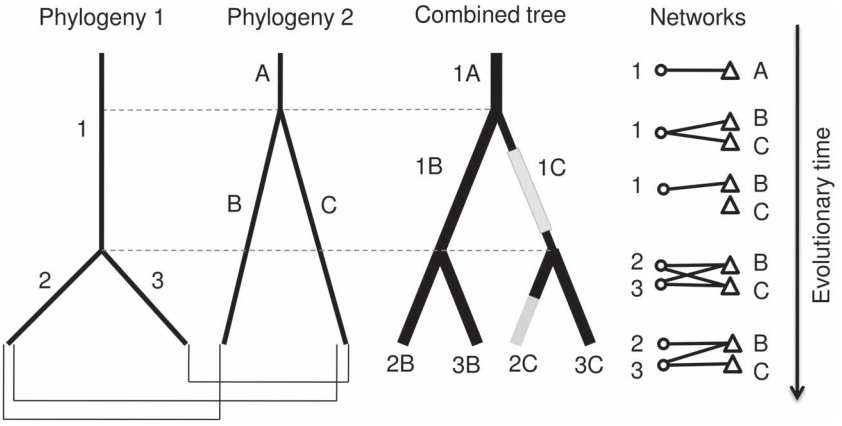

Figure 1. Example of the combination of two phylogenetic trees: every combination of branches of the two phylogenetic trees is represented by a corresponding branch in the combined tree. The branch in the combined tree is coloured grey when the interaction between the corresponding species pair is lost. The right panel shows resulting interaction networks.

parameter K, assigning the branch lengths in the first and second tree respectively as $L_i = (d_i/D)^{\lambda-1}$ and $L_i = K(d_i/D)^{\lambda-1}$ where L_i is the length of the branch leading to node i, d_i the depth of node i (i.e. the number of branches separating node i from the root) and D the height of the tree (i.e. the depth of the deepest leaf node) – Supplementary material Appendix 1 Fig. A1. This results in all branch lengths being equal when $\lambda = 1$, with branch lengths increasing/decreasing from the root to the leaves when $\lambda > 1$ and $\lambda < 1$ respectively. Of course, when dated phylogenies are available, they should be used to replace the above proxy in the Markov process model.

Network evolution on the combined tree

To model the evolution of an entire network through time, we associate each branch of the combined tree (corresponding to the time during which a particular pair of species co-existed) with a Markov process describing the evolution of the interaction between the pair of species in question. We assume that, when a branch in the combined tree splits (due to a speciation event in one of the corresponding species), the descendant species pairs inherit the interaction state (present or absent) of the parent species pair. Using this model, and given the two species phylogenies and the model parameters μ, π_0, λ and K, the likelihood of an observed interaction network can then be calculated using Felsenstein's pruning algorithm (Felsenstein 1981). We estimated π_0 as the empirical frequency of s_0 for each observed network, while the remaining parameters were inferred using maximum likelihood. The model and the inference of μ and π_0 were implemented as scripts for the HyPhy software package (Kosakovsky Pond et al. 2005) and are available at <https://github.com/spond/pubs/>. The parameters λ and K affect the structure of the combined tree and could therefore not be optimized in HyPhy; instead, we optimized them by calculating the likelihood for all values in a pre-specified range (Table 1) and selected the maximum likelihood ones as the optimized combined tree.

Data sets

For empirical ecological networks, we chose the 53 interactive communities compiled by Rezende et al. (2007b; pers. comm.), composed of 31 plant–pollinator and 22 plant–frugivore mutualistic networks (Supplementary material Appendix 1 Table A1). Each network was represented by its adjacency matrix A with rooted phylogenies of the animals and plants in the known network. The model took both the adjacency matrix and the two phylogenies as input. Branch lengths were available only for 18 animal communities. Using these known branch lengths for animals but not plants led to a poor model fit, possibly due to inaccurate relative calibration of the two trees while constructing the combined tree, so we only present results using our branch length approximation for analyses.

Table 1. Summary of model parameters.

Symbol	Definition	Full range studied
μ	gain–loss rate parameter	$[0, +\infty)$
π	vector of equilibrium frequency: $\pi = (\pi_0, \pi_1)$ and $\pi_0 + \pi_1 = 1$	π_0 and π_1 [0,1]
λ	parameter controlling the variation of branch lengths along the tree	given by 2^n for n $[-5,-4,\ldots,4, 5]$[1]
K	parameter which scales the height of the two phylogenetic trees relative to each other	[0.25, 0.5,…,3.25, 3.5][2]

[1]beyond this range, branch lengths are numerically indistinguishable from zero; [2]beyond this range, we obtain overly long branches, leading to saturation.

Testing for phylogenetic signals

To evaluate the importance of the evolutionary history in shaping the observed networks (i.e. the strength of phylogenetic signals in the interaction networks), we simulated a set of 100 pairs of random phylogenies for each network and optimized the model parameters in the same way as for the observed networks. This allowed us to calculate, for each network, the null distribution of maximum likelihood scores. If the score of the correct phylogeny was in the upper 5% of the simulated null distribution, we rejected the null hypothesis that the phylogeny is uninformative.

Evaluating the node degree distribution

To evaluate the distributions of node degree predicted by our model, we compared it to four widely reported parametric models of degree distribution (Jordano et al. 2003): the power law, truncated power law, discretized exponential, and negative binomial distributions, defined on positive integer degrees. We also compared with a null model (Null model I in Bascompte et al. 2003) in which the probability of the presence of an interaction between species i and j is equal to the connectance (i.e. the proportion of 1s in the matrix). From those matrices, we constructed a multinomial model of node degree distribution in the same way. The above parametric models can be used to assign likelihood scores and optimize model fit to the observed node degree distributions; however, unlike our phylogenetic likelihood model, these curve-fitting models cannot fine-tune ecological and evolutionary processes in order to optimize fit to the observed networks themselves.

To facilitate comparison with these parametric models of node degree distribution, we adapted our model to optimize the fit to the observed node degree distribution rather than the full network. For each set of parameter values, we used our model to simulate 1000 interaction networks. By adding the degree counts of individual networks, we formed a histogram representing a single estimate of the node degree distribution for each parameter setting. We smoothed this estimate by adding a pseudocount of one to all degree values (Durbin et al. 1999) – this avoids point estimates of 0 frequencies for any node degree between 1 and the maximum observed value. We interpreted the resulting degree distribution as a probability density function which specifies, for any node, the probability with which our model predicts it to have a given degree. Assuming mutually independent node degrees, we calculated a likelihood value for the observed node degrees in

every real network and for every combination of model parameters – this is a commonly used approach for analyzing degree distributions (Handcock and Jones 2004). Finally, since the model parameters λ and K were originally optimized by brute force calculation on a grid of values, we were able to re-optimize them for each network by selecting those values which maximized the likelihood of the observed degree distribution (this option was not available for μ, which was originally optimized via standard optimization techniques, or π_0, which is set to the empirical frequency throughout; reusing parameter estimates obtained from a simpler version of the model is a standard computational shortcut (Murrell et al. 2012) and any resulting suboptimality would only decrease the performance of the model without invalidating our claims).

Given a network, alternative models can be fitted to its degree distribution using maximum likelihood, and their goodness of fit evaluated using the small sample correction of the Akaike information criterion (AIC_c; Burnham and Anderson 2002). We computed the ΔAIC_c scores, which are the AIC_c scores for each model minus the AIC_c score for the best model on the data set in question. Thus $\Delta AIC_c = 0$ for the best-supported model while, as a rule of thumb (Burnham and Anderson 2002), models with $\Delta AIC_c \leq 2$ have substantial empirical support, models with larger ΔAIC_c have considerably less and models with $\Delta AIC_c > 10$ have no support.

To evaluate how well our model generally predicts the degree distribution, we also conducted the Kolmogorov–Smirnov test for comparing the degree distributions predicted by the model (given by parameters maximizing the full network) with the observed ones. To assess how much of the variation in observed degrees can be explained by the model we compared, for each network, the predicted and observed frequencies of each node degree using a reduced major axis (RMA) linear regression (Bohonak and van der Linde 2004).

Evaluating the predicted nestedness

The nestedness of observed and predicted networks was measured by the NODF (nested metric based on overlap and decreasing fill) (Almeida Neto et al. 2008) using the Aninhado software (Guimarães and Guimarães 2006). This measure quantifies the degree of nestedness of a network, ranging from 0 (no nestedness) to 100 (completely nested). We calculated, for each network, the NODF of the empirical network and those of 1000 simulations from our model so that a p-value could be derived. We further simulated 1000 networks from a null model (null model II in Bascompte et al. 2003) with the probability of an interaction between species i and j given by $p = 0.5(N_i/C + N_j/R)$, where N_i (resp. N_j) is the number of 1s in row i (resp. in column j), C and R the number of columns and rows. This null model can produce approximately the same shape of degree distribution but lacks the nested structure. We assessed the fit of both our model and the null model to the empirical data by means of a two-tailed hypothesis test: if the observed NODF was in the upper or lower 2.5% of the simulated distribution, the corresponding model was rejected. To further quantify the relationships between the observed NODF and the NODFs predicted by our model and by the null model, we plotted the predicted NODF versus the observed ones (Fig. 4) and used the RMA (reduced major axis) method (Bohonak and van der Linde 2004) to perform a linear regression.

To allow comparison between networks and to diminish the effect of network sizes on the absolute NODF, we also calculated the relative NODF given by $N^* = (N - N_r)/N_r$, in which N is the absolute NODF and N_r is the average value of nestedness for replicates of our model or the null model (Bascompte et al. 2003). As some of the results considered significant (with a p-value ≤ 0.05) could be false positives, we also present results in which the proportion of false positives is controlled by fixing the false discovery rate (FDR) (Benjamini and Hochberg 1995), which is the expected percentage of false positives among all the significant hypotheses.

Evaluating the predicted modularity

The modularity of observed and predicted networks was measured using the software NETCARTO based on simulated annealing for the modularity optimisation (Guimera and Amaral 2005). NETCARTO quantifies the modularity of a network ranging from 0 (random) to 1 (completely compartmentalised). For each network, we computed the modularity for the 1000 replicates of our model as well as for the observed network. As in the nestedness analysis, we assessed the fit of our model to empirical data by means of a two-tailed hypothesis test in which our model was rejected when the empirical modularity value lies outside the 95% confidence interval of the modularity distribution from our model.

Results

Phylogenetic signal

For 11 of the 53 empirical networks (21%), we were able to reject (at $p = 0.05$) the null hypothesis that the network is not informed by the phylogeny (see two examples in Fig. 2). For the majority of networks, however, we were unable to reject the null hypothesis: these networks could also have arisen under different evolutionary histories. The average estimated number of interaction gain/loss events per network (i.e. the total over the entire history of the network) was 2.6 in cases where the null hypothesis was rejected and 4.8 in cases where it was not. The difference was not statistically significant (one-tailed t-test: $p = 0.074$), but is nonetheless consistent with our expectation that frequent gain/loss events from ecological/behavioural change will erase the effect of phylogenetic structure especially in networks with a deeper history (i.e. longer branches).

Some typical interaction networks sampled from the predictive distribution of our model are shown alongside the observations from real networks in Fig. 3 and the Supplementary material Appendix 1 Fig. A2. On average, 55.2% of the variance of node degree frequencies was explained by our phylogenetic model according to the RMA regression (Supplementary material Appendix 1 Table A3, Fig. A3).

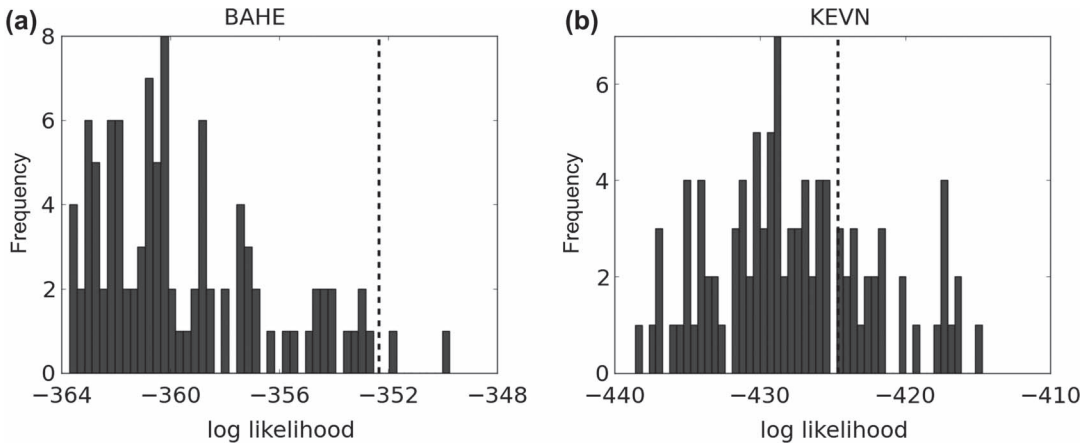

Figure 2. Likelihood histogram of the analysis of two different networks: BAHE and KEVN. Likelihood values when using random phylogenies are represented by the closed bars while the broken line indicates the value of the likelihood when using the correct phylogeny. For the BAHE network we cannot rule out the hypothesis that the network is independent of phylogeny, while for KEVN we can.

Network architecture

The AIC_c analysis revealed that the node degree distributions of 34% (18/53) of the networks were best fit by our model and 36% (19/53) by the truncated power law (Table 2; see Fig. 4 for two examples). Therefore, although in the majority of cases our model fails to predict the node degree distribution as well as specialised parametric models, it outperforms these models on a third of the networks investigated. The Kolmogorov–Smirnov test showed that for 43% of the networks, the differences between predicted and observed degree distributions were not significant (Supplementary material Appendix 1 Table A2). This finding is encouraging, because our model was not explicitly set up to match the observed distribution of node degrees: the phylogenies of the species alone appear to be sufficient to recapitulate it. Furthermore, when the node degree frequencies predicted by our model were plotted against the observed frequencies (Supplementary material Appendix 1 Fig. A3), in 53% of the networks more than half of the variation was explained by our model ($R^2 > 0.5$). This suggests that, for most networks, empirical degree distributions and simulated degree distributions have a strong positive correlation (Supplementary material Appendix 1 Table A3).

Of the 35 networks best fit by the parametric models of degree distribution, 37% are frugivory networks, whilst 50% of the 18 networks best fit by our model are frugivory networks. This suggests that compared to other models, our model could be more suitable for frugivory networks rather than pollination networks. However, Fisher's exact test did not reveal a significant association ($p = 0.394$) between network type and whether the network is well fit by our model. Moreover, networks best fit by our model have a significantly smaller size (with an average number of 53.4 species) than networks best fit by other models (average number of 99.6 species) (*t*-test, $p = 0.014$). Networks best fit by our model tended to be more nested (with a mean of the nestedness values: 43.0) than networks best fit by other models (mean of the nestedness values: 33.4), although this was not significant (*t*-test, $p = 0.09$).

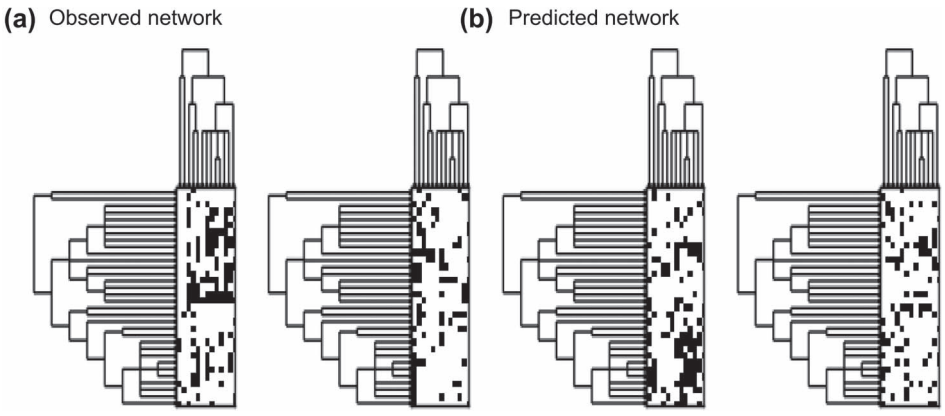

Figure 3. Empirically observed interaction network (a) of a pollination community (reference: SMAL, Supplementary material Appendix 1 Table A1) with three realizations of the same network by our model (b). Animal species (on the rows) and plant species (on the columns) are given with their respective phylogenetic trees. Similarities between the empirical and simulated networks are indicative of the extent to which phylogenetic history influences the network structure.

Table 2. Number of networks fit by each candidate model for node-degree distribution.

	Best model $\Delta AIC_c = 0$	Strongly supported $0 < \Delta AIC_c \leq 2$	Weakly supported $2 < \Delta AIC_c \leq 10$	No support $\Delta AIC_c > 10$
Truncated power law	19	8	13	13
Markov process model	18	2	13	20
Power law	7	6	9	31
Negative binomial	5	7	14	27
Exponential	3	6	12	32
Null model	1	1	10	41

Based on observed nestedness, null model II was rejected (at a $p = 0.05$) for 43 of the 53 networks (81%; see Supplementary material Appendix 1 Table A4 for details), suggesting the majority of the empirical networks were significantly nested. Under the same test, our model was rejected for 24 networks (45%), producing nestedness predictions that could not be rejected for the remaining 55% of the real networks. When the FDR is controlled to be lower than 0.05, null model II was rejected for 79.2% of the networks, and our model was rejected for 41.5% of the networks, consistent with the above results. We also obtained similar results when using the relative NODF (Supplementary material Appendix 1 Table A5). When the average values of the predicted NODF of simulated networks were plotted against the observed values (Fig. 5), our model performed better for higher degrees of nestedness, compared to the null model. Taken together, these results support the claim that nested structure matching that observed in real networks emerges naturally in our model.

For the modularity comparison, the two-tailed test showed that for 48 of the 53 networks (90%), our model predictions are not significantly different from the observed modularities – this makes our model a strong candidate for explaining the emergence of compartmentalisation in networks (Supplementary material Appendix 1 Table A6). When the average values of modularity of simulated networks were plotted against the observed modularity values (Fig. 6), the RMA regression showed a coefficient of determination of 0.69, indicating a linear relationship and nearly 70% variation of observed modularity explained by our model.

Interaction heritability

Values of the parameter λ (see Table 1 for the definition of each parameter) that maximised the likelihood of the data were all less than one, suggesting an increase of branch lengths from the root to the leaves. Because the optimized branch lengths were inferred from the observed interaction matrices, the long branches at leaf nodes could reflect a high level of ecological dissimilarity between closely related species, as compared to a larger degree of ecological similarity between related clades. No systematic pattern was identified for the parameter *K,* which controls the relative scales of the trees.

Estimates of the gain-loss rate parameter (μ) ranged from 0.14 to 6.16 with an average of 1.34 changes per unit of branch length. For most of the networks (47 out of 53), the

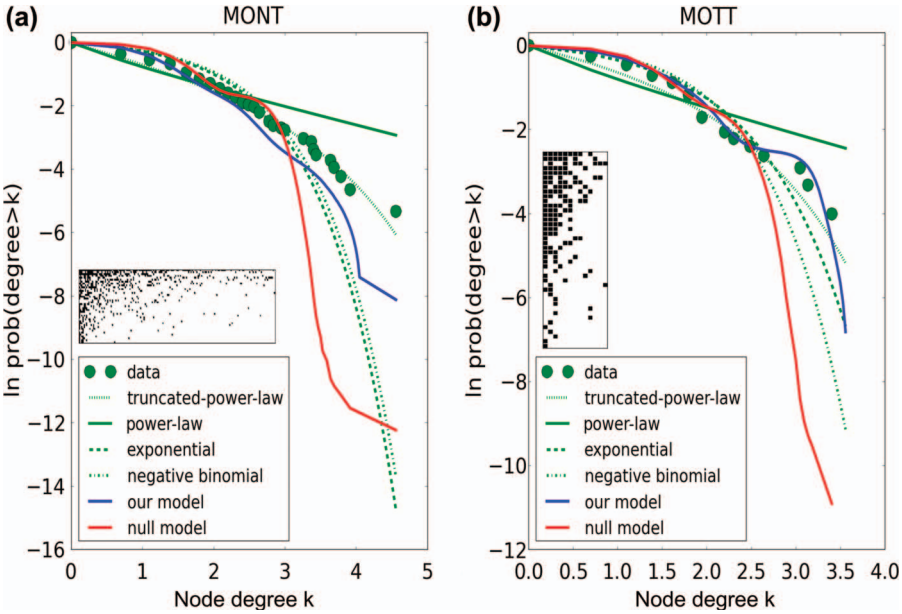

Figure 4. Log-log plot of the complementary cumulative distribution of the connectivity of a frugivory network (MONT) and a pollination network (MOTT). Inset: adjacency matrix representation. The degree distribution in MONT is best fit by a truncated power-law while it is best fit by the Markov process model in MOTT.

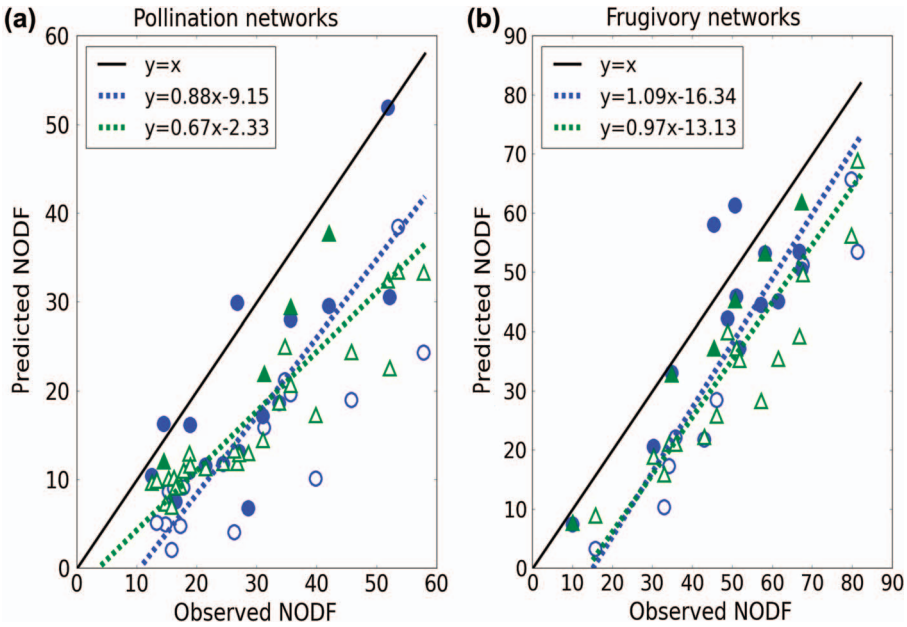

Figure 5. Average NODF values of pollination (a) and frugivory (b) networks simulated under our model or under the null model vs. NODF values of empirical networks. Continuous lines represent $y = x$. Blue circles represent our model predictions (fit by blue dashed lines) and green triangles represent null model predictions (fit by green dotted lines). The triangle or the circle is empty when the candidate model is rejected (p-value < 0.05). Otherwise, the network is represented by a solid triangle or circle.

existence of an interaction between two species appears to be less probable than the absence of an interaction ($\pi_0 > \pi_1$, where π_0 and π_1 are the equilibrium frequency vector). Taking into account the amount of evolutionary time since the most recent common ancestors of the species existed (given by the total length of the phylogenetic trees), ecological changes (gain and loss events) are estimated to have occurred on average 3096 times per network or 4.12 times per path from the root to a leaf node.

Discussion

We have presented a model that uses phylogenetic history to describe not only the structure of extant ecological interaction networks, but also the emergence of these structures. Despite not having reliable branch length estimates available, we found that the phylogenetic structure was informative in at least a substantial minority (21%) of networks. Our model is of interest for two principal reasons.

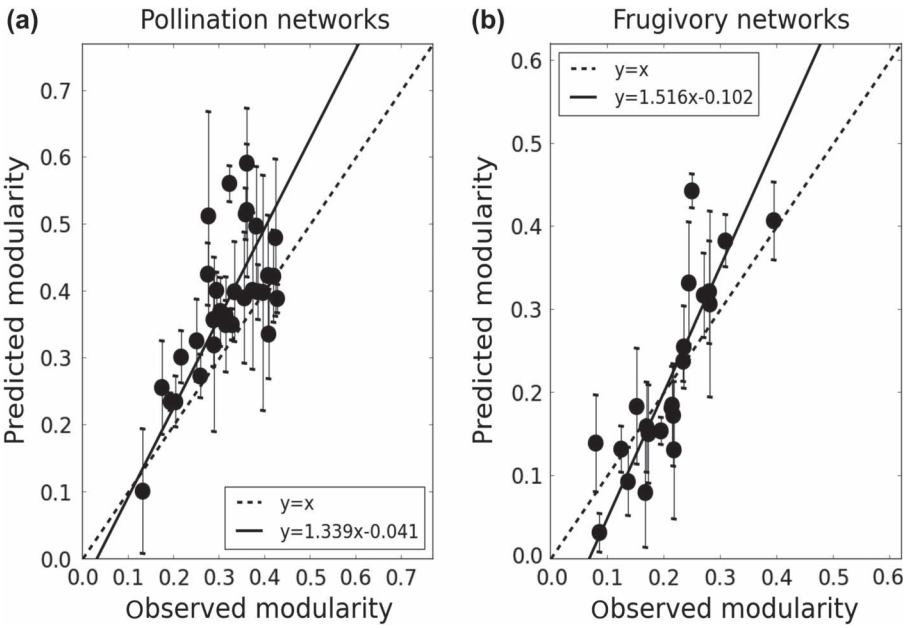

Figure 6. Average modularity values of (a) pollination and (b) frugivory networks simulated under our model versus modularity values of empirical networks. Error bars indicate one standard deviation.

First, existing models typically focus on isolated aspects of network structure (Rezende et al. 2007a, b, Takemoto and Arita 2010), such as node degree distribution and nestedness, without necessarily considering these metrics simultaneously and while ignoring network aspects not summarised by them. By contrast, our model predicts complete network structures along with putative evolutionary network histories. Second, we found that even our simple generative description provides qualitatively realistic predictions of the resulting network structure, in many cases even quantitatively outperforming special purpose parametric models of node degree distribution and nestedness.

The previous study by Rezende et al. (2007a, 2007b) on the importance of coevolution to network structure is based on statistical analysis of the correlation between observed network structure and phylogenetic history. By contrast, our approach was able to partially reproduce the network structures, along with putative histories describing their formation, as emerging directly from a dynamical model informed by phylogeny. This supports (in a subset of networks) the finding that detectable phylogenetic signals may be typical of mutualistic networks. The fact that observed levels of modularity are well reproduced by our model supports the argument that modularity results largely from coevolution (Olesen et al. 2007). Moreover, the poor fit of our model to observed nestedness suggests that nested structures in networks likely reflect ecological constraints from the spatiotemporal overlap of partner species (Morales and Vázquez 2008, Encinas-Viso et al. 2012) and behavioural mechanisms (Zhang et al. 2011), while coevolutionary processes disrupt nested structures (Nuismer et al. 2013).

Scientists routinely assume that the factors determining the interaction partners of individuals are shared among different individuals of the same species – without this assumption, it would make no sense to refer to interactions between 'species' rather than individuals. As a consequence, individual members of a species are assumed to share the interaction partners of their parents, although interactions (of both the parents and their offspring) can be gained or lost over time. Our model assumes nothing stronger than this: just as genetic characteristics are inherited from one generation to the next, so are interaction partners. Such inheritance could be due to inheritance of genetic characteristics (e.g. morphology for obtaining nectar from a particular flower, or digestive physiology and the ability to process a particular toxin, which together potentially constrains feeding on a particular resource) or purely for reasons of spatial proximity (to the extent that individuals of subsequent cohorts occur within the same location and environment as their parents, they are likely to be exposed to similar availability of resources in the form of potential interaction partners). No differences in the mode of inheritance are assumed during speciation: we conceive of speciation as a gradual process occurring over many generations, with inheritance of both genetic and interaction characteristics happening at every generation; when subpopulations become reproductively or spatially separated, the ongoing processes of inheritance within the subpopulations remain unaffected. Thus we see the key difference between the evolution of genetic and interaction characteristics as lying not in inheritance from one generation to the next, but in the time scales over which those characteristics are conserved.

For phylogenetic structure to be informative regarding the emergence of interaction networks, it is required that interaction characteristics are conserved over evolutionary time scales. When this is not the case, ecological change is fast compared to evolutionary change and our model is expected to infer long branch lengths, effectively reverting to a model that is uninformed by phylogenetic structure and which we do not expect to fit the data any better than one based on random phylogenies. However, in many cases we do expect interaction patterns to be conserved because interactions between species depend on the complementarity of phenotypic traits of the species (Santamaría and Rodríguez-Gironés 2007, Anderson et al. 2010).

Dealing with the right phylogeny and an accurate set of branch length estimates is crucial. The unavailability of pre-estimated branch lengths for our phylogenies undoubtedly had a negative impact on the fit of our model. Our branch length estimates are almost certainly unreliable, as they were based on the presence or absence of ecological interactions rather than on genetic or morphological features. For this reason, it is perhaps remarkable that qualitatively and even quantitatively realistic simulations of network evolution could be obtained at all. Our branch length estimation scheme does allow the model to set the internal branch lengths so that they are negligibly short compared to the external branch lengths, thus removing the effect of the phylogenetic structure (the tree becomes a star phylogeny). This means that the null model in the test for phylogenetic signal is unaffected by the unavailability of branch lengths, while the alternative model may be severely hampered – our test is likely to have underestimated the proportion of networks with detectable phylogenetic signal. We expect that the use of less crude branch length estimates, as are increasingly commonly available from genetic sequence data, would further improve the realism and fit of the model.

It is informative to consider the specific role of branch lengths in our model. In species trees, branch lengths can be considered as proportional to time (measured in units of years or generations) or to time multiplied by the rate of evolutionary change. In our model, branch length (after being scaled by the interaction gain/loss rate parameter μ) is proportional to time multiplied by the rate of 'ecological' change, where the changes in question are gains and losses of ecological interactions and the rate of ecological change is modelled as constant across the phylogeny. Short branches imply a relatively small number of ecological network changes between speciation events, which leads to network structure that is strongly affected by phylogenetic structure and hence by evolutionary history. By contrast, long branches generate network structures that are relatively independent of evolutionary history.

Although we investigated only mutualistic networks, our model can be used to describe any bipartite network, including antagonistic networks such as predator–prey and host–parasite networks. We expect that the role of evolutionary history in shaping network structure could be demonstrated in the same way in such examples. Furthermore, extension of our model to cases where interaction

strength is quantified, for instance corresponding to the number of visits or amount of pollen transferred between species (Vázquez et al. 2005), should also be feasible. This could be done by using weighted graphs instead of our binary presence/absence representation (Bascompte et al. 2006) and extending the Markov process to use a continuous state space.

Our description of network evolution is extremely simple, and it is therefore perhaps remarkable that the aspects of network structure investigated here can often be explained by phylogenetic history alone. Among the mechanisms neglected by our model are selective effects, which may favour the formation of beneficial interactions. Evolutionary constraints may also force interacting species to remain associated despite lower average fitness (Zhang et al. 2011) than would be the case if interacting species changed their partners. Another simplification was the assumption of rate homogeneity: our model does not allow the rates at which interactions are gained and lost to change over time or across the phylogeny, as might be expected in real biological systems. A next step in developing this work would be to develop models with increasing degrees of biological realism. For instance, we have recently developed methodology by which the rate homogeneity assumption can be relaxed (Kosakovsky Pond et al. 2011, Murrell et al. 2012).

We see our model as providing an initial platform on which more sophisticated generative models can be built, incorporating factors such as trait selection, species rewiring (interaction switching) and interspecific competition, which are no doubt of great importance in shaping the structure of interactive networks. Ultimately, an aim of this approach is not only to understand the current structure of observed interaction networks, but also the evolutionary and ecological processes leading to their emergence. To this end, our model may be used as a null model (especially for controlling the phylogenetic signal) when assessing the evidence for and the relative importance of many potentially contributing factors.

Acknowledgements – We are grateful to Jordi Bascompte for providing the networks with phylogenies. This project was supported by the African Inst. for Mathematical Sciences (AIMS), the South African DST-NRF Centre of Excellence for Invasion Biology, NIGMS R01 award 1R01GM093939 and NIH awards DA034978 and AI036214. CH., JST and KS received support from the South African NRF Incentive Funding Programme and CH is supported by the South African Research Chair initiative (SARChI). We thank the associate editor for constructive comments on the manuscript.

References

Almeida Neto, M. et al. 2008. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. – Oikos 117: 1227–1239.

Anderson, B. et al. 2010. Predictable patterns of trait mismatches between interacting plants and insects. – BMC Evol. Biol. 10: 204.

Barabási, A. L. and Albert, R. 1999. Emergence of scaling in random networks. – Science 286: 509–512.

Bascompte, J. and Jordano, P. 2007. Plant–animal mutualistic networks: the architecture of biodiversity. – Annu. Rev. Ecol. Evol. Syst. 38: 567–593.

Bascompte, J. et al. 2003. The nested assembly of plant–animal mutualistic networks. – Proc. Natl Acad. Sci. USA 100: 9383–9387.

Bascompte, J. et al. 2006. Asymmetric coevolutionary networks facilitate biodiversity maintenance. – Science 312: 431–433.

Benjamini, Y. and Hochberg, Y. 1995, Controlling the false discovery rate: a practical and powerful approach to multiple testing. – J. R. Stat. Soc. 57: 289–300.

Bohonak, A. J. and van der Linde, K. 2004. RMA: Software for reduced major axis regression, Java version. – <www.kimvdlinde.com/professional/rma.html>.

Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference. – Springer.

Churchill, G. A. 1989. Stochastic models for heterogeneous DNA sequences. – Bull. Math. Biol. 51: 79–94.

Dicks, L. V. et al. 2002. Compartmentalization in plant–insect flower visitor webs. – J. Anim. Ecol. 71: 32–43.

Durbin, R. et al. 1999. Biological sequence analysis: probabilistic models of proteins and nucleic acids. – Cambridge Univ. Press.

Encinas-Viso, F. et al. 2012. Phenology drives mutualistic network structure and diversity. – Ecol. Lett. 14: 198-208.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. – J. Mol. Evol. 17: 368–376.

Felsenstein, J. 2004. Inferring phylogenies. – Sinauer.

Grafen, A. 1989. The phylogenetic regression. – Phil. Trans. R. Soc. B 326: 119–157.

Guimarães Jr., P. R. and Guimarães, P. 2006. Improving the analyses of nestedness for large sets of matrices. – Environ. Model. Software 21: 1512–1513.

Guimarães Jr., P. R. et al. 2007. Build-up mechanisms determining the topology of mutualistic networks. – J. Theor. Biol. 249: 181–189.

Guimarães Jr., P. R. et al. 2011. Evolution and coevolution in mutualistic networks. – Ecol. Lett. 14: 877–885.

Guimera, R., and Amaral, L. A. N. 2005. Cartography of complex networks: modules and universal roles. – J. Stat. Mech. Theory Exp. P02001.372.

Handcock, M. S. and Jones, J. H. 2004. Likelihood-based inference for stochastic models of sexual network formation. – Theor. Popul. Biol. 65: 413–422.

Holland, J. N. and DeAngelis, D. L. 2010. A consumer–resource approach to the density-dependent population dynamics of mutualism. – Ecology 91: 1286–1295.

Hui, C. et al. 2013. Increasing functional modularity with residence time in the co-distribution of native and introduced vascular plants. – Nat. Commun. 4: 2454.

Jordano, P. et al. 2003. Invariant properties in coevolutionary networks of plant animal interactions. – Ecol. Lett. 6: 69–81.

Kaiser-Bunbury, C. N. et al. 2010. The robustness of pollination networks to the loss of species and interactions: a quantitative approach incorporating pollinator behaviour. – Ecol. Lett. 13: 442–452.

Kosakovsky Pond, S. L. et al. 2005. HyPhy: hypothesis testing using phylogenies. – Bioinformatics 21: 676–679.

Kosakovsky Pond, S. L. et al. 2011. A random effects branch-site model for detecting episodic diversifying selection. – Mol. Biol. Evol. 28: 3033–3043.

Krishna, A. et al. 2008. A neutral-niche theory of nestedness in mutualistic networks. – Oikos 117: 1609–1618.

Lemey, P. et al. (eds) 2009. The phylogenetic handbook, 2d edn. – Cambridge Univ. Press.

Medan, D. et al. 2007. Analysis and assembling of network structure in mutualistic systems. – J. Theor. Biol. 246: 510–521.

Morales, J. M. and Vázquez, D. P. 2008. The effect of space in plant–animal mutualistic networks: insights from a simulation study. – Oikos 117: 1362–1370.
Murrell, B. et al. 2012. Detecting individual sites subject to episodic diversifying selection. – PloS Genetics 8: e1002764.
Nuismer, S. L. et al. 1999. Gene flow and geographically structured coevolution. – Proc. R. Soc. B 266: 605–609.
Nuismer, S. L. et al. 2013. Coevolution and the architecture of mutualistic networks. – Evolution 67: 338–354.
Okuyama, T. 2008. Do mutualistic networks follow power distributions? – Ecol. Complexity 5: 59–65.
Olesen, J. M. et al. 2007. The modularity of pollination networks. – Proc. Natl Acad. Sci. USA 104: 19891–19896.
Olesen, J. M. et al. 2008. Temporal dynamics in a pollination network. – Ecology 89: 1573–1582.
Rezende, E. L. et al. 2007a. Effects of phenotypic complementarity and phylogeny on the nested structure of mutualistic networks. – Oikos 116: 1919–1929.
Rezende, E. L. et al. 2007b. Non-random coextinctions in phylogenetically structured mutualistic networks. – Nature 448: 925–928.
Rodríguez-Gironés, M. A. and Llandres, A. L. 2008. Resource competition triggers the coevolution of long tongues and deep corolla tubes. – PLoS ONE 3: e2992.
Santamaría, L. and Rodríguez-Gironés, M A. 2007. Linkage rules for plant–pollinator networks: trait complementarity or exploitation barriers? – PLoS Biol. 5: e31.
Stroock, D. W. 2005. An introduction to Markov processes. – Springer.
Suweis, S. et al. 2013. Emergence of structural and dynamical properties of ecological mutualistic networks. – Nature 500: 449–452.
Takemoto, K. and Arita, M. 2010. Nested structure acquired through simple evolutionary process. – J. Theor. Biol. 264: 782–786.
Vázquez, D. P. 2005. Degree distribution in plant–animal mutualistic networks: forbidden links or random interactions? – Oikos 108: 421–426.
Vázquez, D. P. and Aizen, M. A. 2003. Null model analyses of specialization in plant–pollinator interactions. – Ecol. Lett. 84: 2493-2501.
Vázquez, D. P. et al. 2005. Interaction frequency as a surrogate for the total effect of animal mutualists on plants. – Ecol. Lett. 8: 1088–1094.
Vázquez, D. P. et al. 2007. Species abundance and asymmetric interaction strength in ecological networks. – Oikos 116: 1120–1127.
Vázquez, D. P. et al. 2009. Evaluating multiple determinants of the structure of plant–animal mutualistic networks – Ecology 90: 2039–2046.
Waser, N. M. et al. 1996. Generalization in pollination system and why it matters. – Ecology 77: 1043–1060.
Yang, Z. 2006. Computational molecular evolution. – Oxford Univ. Press.
Zhang, F. et al. 2011. An interaction switch predicts the nested architecture of mutualistic networks. – Ecol. Lett. 14: 797–803.
Zhang, F. et al. 2013. Adaptive divergence in Darwin's race: how coevolution can generate trait diversity in a pollination system. – Evolution 67: 548–560.

Supplementary material (available online as Appendix oik-00857 at <www.oikosjournal.org/readers/appendix>). Appendix 1 containing: Table A1: Data sets used. Table A2: Kolmogorov–Smirnov comparison of node degree distributions predicted by the model vs empirical observations. Table A3: Statistics of the reduced major axis regression of predicted and observed species probability degree. Table A4: Average absolute NODF values. Table A5: Average relative NODF values. Table A6: Average modularity values. Figure A1: Schematic diagram showing the assignment of branch lengths to the phylogenetic trees in our simulations. (a) Branch lengths are all equal to one. The depth of each node is given by d_i. (b) Branch lengths l_i are assigned as function of d_i and increase from root to leaf nodes. (c) Branch lengths decrease from root to leaf nodes. Figure A2: Presence–absence matrix representations of the 53 networks together with their respective species phylogenies. The empirical network is on the left side of each plot, while three predictions by the model are on the right within each box. Network codes and type of network are given in Table A1. Figure A3: Node degree frequencies predicted by the evolutionary model versus observed in the 53 networks.